DOCUMENT RESUME

ED 025 080 EC 003 039

By- Eber, Herbert W.

Personality Measurement with Mentally Retarded and Other Sub-Cultural Adults. Final Report.

Alabama Inst. for Deaf and Blind, Talladega.
Spons Agency- Vocational Rehabilitation Administration (DHEW), Washington, D.C.
Pub Date 1 Jul 65
Grant- VRA-RD-1059p-62
Note- 43p.
EDRS Price MF-$0.25 HC-$2.35
Descriptors- Cultural Disadvantagement, *Disadvantaged Youth, *Exceptional Child Research, Factor Analysis, Functional Illiteracy, Individual Characteristics, *Mentally Handicapped, Personal Adjustment, *Personality, Personality Assessment, Personality Tests, Predictive Validity, Retarded Readers, Test Construction, *Tests, Vocational Adjustment, Vocational Counseling, Vocational Rehabilitation

Two 160-item experimental forms of multidimensional personality test to assess vocational potential of clients of limited literacy (third grade reading level) were developed and administered to clients at rehabilitation centers and at centers for the retarded. Using the 16 Personality Factors Test as a model, items were constructed to do the following: yield meaningful self-descriptions and personality description patterns that match known ones; determine if matching patterns are sufficiently similar to permit the transfer of previous knowledge to the use of new patterns; and establish the stability of new found patterns. Data analysis consisted of the oblique simple structure criterion and analysis by item packages. Results with 325 clients, whose literacy levels and verbal sophistication were below the sixth grade level (including 216 who completed both forms), implied that items meeting the above standard were derived. Eight analyses revealed the factors which indicated whether a person is outgoing, emotionally stable, impulsive, assertive, enthusiastic, conscientious, venturesome, tender minded, aloof, suspicious, imaginative, shrewd, apprehensive, self sufficient, self disciplined, or tense. (LE)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED025080

FINAL REPORT

Project Number: RD-1059p-62

Project Title: PERSONALITY MEASUREMENT WITH MENTALLY RETARDED AND OTHER SUB-CULTURAL ADULTS

Grant Period Dates: March 1, 1962 through February 29, 1965

Report Written By: Herbert W. Eber, Ph.D.
Research Director
P. O. Box 3361-A
Birmingham, Alabama 35205

Grantee Organization: Alabama Institute for Deaf and Blind
P. O. Box 268
Talladega, Alabama 35160

Date of Preparation: July 1, 1965

This investigation was supported, in part, by a research grant, number RD-1059p-62, from the Vocational Rehabilitation Administration, Department of Health, Education, and Welfare, Washington, D. C. 20201.

ED025080

Table of Contents

I. INTRODUCTION 3
Limited Literacy as a Barrier in Evaluation. 3
Personality Data as Vocational Predictors. 4
II. PROJECT RATIONALE. 4
Purpose. 4
The Test Model 5
Prevalence of the Target Population. 8
Feasibility of the Needed Instrument 9
Relevance of a 16 PF to Vocational Considerations. 10
Appropriateness of Research Setting. 13
III. PROJECT METHOD 14
Factor Analysis. 14
The Client Sample. 15
The Experimental Items 16
Data Collection. 18
IV. DATA ANALYSIS. 20
The Oblique Simple Structure Criterion 20
Analysis by Item Packages. 22
V. RESULTS. 24
Item Packages from Form X. 24
Analysis of Single Items from Form X 27
Analysis by New Form X Packages. 29
Analysis of Form X Packages Combined with Form Y Items 30
Difficulties Inherent in the Package Marker Approach 30
The Large Matrix 32
VI. DISCUSSION 33
Table 1. 37
References 38

A PERSONALITY TEST FOR PERSONS OF LIMITED READING SKILLS

I. INTRODUCTION

Limited Literacy as a Barrier in Evaluation

In recent years, vocational rehabilitation has turned increasing attention to the problems of mentally retarded clients. Even prior to the emphasis upon this handicapping condition during the administration of President Kennedy, a number of projects in the area of service to the mentally retarded were under way. With increased emphasis upon such service by all branches of the federal and state governments, the primary need for vocationally oriented services became more and more recognized.

In parallel with increased attention to mentally retarded clients, there also appeared increased recognition of the relative prevalence of culturally deprived persons in the vocational rehabilitation client load. This has been particularly true as vocational rehabilitation counselors in the various states have focused more attention upon the so-called difficult clients, the clients with multiple handicaps, and those whose relatively poor education served to make a major vocational handicap out of a relatively minor disabling condition. While poorly educated persons are by no means the same as mentally retarded ones, the two groups do have one thing in common, namely, inability to read test materials (or anything else) at levels involving normal literacy.

Thus there exist, in the population which vocational rehabilitation agencies seek to serve, a large number of clients who cannot read the

most commonly used personality questionnaires. Leaving aside the issues of the validity of such questionnaires for the time being, it certainly seems clear that no validity is possible where the client cannot understand the questions.

Personality Data as Vocational Predictors

Research in a variety of areas has shown the importance of personality characteristics as potential predictors of vocational functioning. Rehabilitation studies conducted by the staff at MacDonald Training Center (1963), as well as other studies, indicated that IQ ceases to be a predictor of vocational adequacy once the range is restricted to those persons usually classified as "mentally retarded". Preliminary studies in a local setting had indicated that personality patterns were likely to be of significant importance in determining the vocational potential of mentally retarded clients. Although unsolved questions of validity still confuse many of the critical issues in this area, there has been growing recognition of the importance of personality characteristics as predictors of vocational functioning among persons in general, without limitation to the mentally retarded or culturally deprived. It certainly appeared reasonable that an instrument suitable for the measurement of important personality characteristics in a group of mentally retarded or culturally deprived persons could prove useful in rehabilitation planning.

II. PROJECT RATIONALE

Purpose

The project was designed to develop data that would lead to a multidimensional personality test suitable for use with clients of limited

literacy. Test items were to be developed and evaluated, factor structure to be identified, and initial data for test construction to be gathered. The third-grade reading level was taken as a target; it was hoped that all required client responses would be such that a "semi-literate" person, one able to read simple prose, could respond meaningfully and effectively.

The project was not directly concerned with test development. Rather, the relevant areas of concern were defined as follows:

(1) Can items yielding meaningful self-descriptions be constructed within the limitations of simplicity and low literacy?

(2) Do these items yield personality description patterns that match known ones?

(3) Are matching patterns sufficiently clearly the same so that validity data on existing patterns can be taken as a useful beginning approximation of predictive inferences from the new items?

(4) Are the new-found patterns sufficiently stable to form the basis of research into their predictive power?

The Test Model

The theoretical and practical model for the project was the 16 Personality Factor Test (Cattell & Eber, 1964), an instrument that attempts to measure relevant personality characteristics through the use of questionnaire items. The test, and the theory behind it, are discussed in substantial detail by Cattell (1957); a substantial bibliography appears in the Handbook (Cattell & Eber, 1964). Only a brief discussion appears appropriate in this report.

The 16 PF Test is a questionnaire designed to measure factors of personality as that term is technically defined. Based upon over twenty years of research by Cattell and his associates, the test is oriented toward the description of those factors which characterize the normal personality, and in this it differs sharply from other available test instruments, which are generally oriented toward psychopathology. This is not to say that psychopathology does not appear on these tests, since substantial evidence indicates the test to be highly sensitive to the identification of neurosis and other forms of maladjustment (Cattell & Scheier, 1961). However, the test was initially developed to measure those factors found appropriate for describing variations among normal personalities, and it seemed reasonable, on a priori and on empirical grounds, that such an instrument would have maximum usefulness as a predictor of vocational functioning.

From the empirical standpoint, a large number of vocational associations of the various factors have been discovered. The Handbook lists profiles, specification equations, and other patterns describing a large variety of occupational groups, and these data not only have been found predictive in a variety of situations but are characterized by an inherent sensibleness that makes the entire theoretical structure quite compelling. It is reasonable, for example, that priests and ministers should be highest in sensitivity among all occupational groups studied, and this is in fact the case; it is reasonable that high-ranking business executives should be persons of superior self-discipline and this is the case; one would expect a group of Olympic championship athletes to be

persons of superb resistance to stress, and this has been the finding; examples could be multiplied, but the reader is referred to the Handbook.

The writer had occasion to begin use of the 16 PF Test in the vocational rehabilitation setting in 1958. Accustomed to the use of general intelligence and projective personality test data, the writer found two areas of inadequacy using these techniques. One was that the cost of evaluation with these techniques tended to be so high that rehabilitation counselors were limited to obtaining psychological evaluation of relatively few of their clients; the other problem was that the type of information developed by these instruments was useful primarily from the standpoint of planning psychotherapy, and was only peripherally useful to the rehabilitation counselor. Adding aptitude measurements to the intelligence-projective test battery made the situation a little better in regard to applicability of the results, but even worse in regard to expense. Dropping the projective tests from the combined battery resulted in little loss from the standpoint of the rehabilitation counselor, but the psychologist could not help but feel that a large area of predictive variance was being ignored.

It is unfortunate, but nevertheless true, that few psychologists in clinical and applied areas had attempted to take seriously, in the applied situation, the laboratory results obtained by Cattell and his co-workers up to about 1958. There appears to be a natural antipathy between the laboratory and the clinic, and in this case it was compounded by the heavy mathematical emphasis of the research that had been published, an emphasis that does not lend itself to good communication with clinically oriented psychologists.

The writer attempted the use of the 16 PF Tests with rehabilitation clients, limiting this use to those clients that were able to read Form C, the simplest form available at that time. No formal research was done at the time, but the results appeared encouraging in the opinion of the counselors with whom the psychologist then consulted. Greater experience in the use of the tests brought apparently even better results, and by 1960 and 1961, the writer was using the 16 PF Test, Form C as a routine part of the rehabilitation battery.

Prevalence of the Target Population

An obvious problem was that the 16 PF, Form C, although it was the simplest of the three forms then available, was still limited in use to those persons able to read at about the sixth-grade level or better. Careful records were kept of the rehabilitation clients referred to the writer, and it was soon discovered that between half and two-thirds of the total client load were unable to deal with the language of Form C.

The writer began use of the Wide Range Achievement Test to categorize incoming clients as (1) "normally literate", e.g., sixth-grade or higher reading level, (2) "semi-literate", e.g., reading level between third and sixth grade, and (3) "functionally illiterate", e.g., reading level below third grade. Over a period of several years, and covering a total of over 3,000 clients, the percentages in each category have remained quite stable at approximately 40% "normally literate", 30% "semi-literate", and 30% "functionally illiterate". It should be emphasized that this does not represent a population of retardates, but rather a general group of clients referred by VR counselors for psychological evaluation.

These data seemed clear in defining the need. If approximately 60% of the referred clients were unable to respond meaningfully to available instruments, the need for a simpler version of these instruments appeared obvious.

Feasibility of the Needed Instrument

Upon reflection, and upon discussion with Cattell, it appeared reasonable to think that a sufficiently simple test measuring the factors could be constructed. In support of such a viewpoint was the fact that success had been encountered in the construction of comparable tests across wide age ranges. Available at that time were three 16 PF Tests, Forms A, B, and C, spanning the educational level from the college level person at the top down to sixth-grade reading level; the High School Personality Questionnaire appropriate for normally literate adolescent clients; and the Child Personality Questionnaire designed for persons between ages 8 and 12, and constructed at about the third-grade reading level. Moreover, the Early School Personality Questionnaire, an instrument which required no reading skills at all for administration, was in process of development and the initial research with it had appeared promising.

Further evidence in support of the expectation that an appropriate test could be constructed came from the fact that the 16 PF had been successfully translated into several languages, had been used with essentially undiminished validity in cross-cultural studies between the English and American populations, and generally had been shown to be relatively "hardy" under conditions of translation and application to persons of quite different cultural backgrounds. It seemed reasonable to assume that a test

which can be translated into Chinese, administered to persons whose native language and culture is Chinese, and yield reasonably comparable measurements of the same personality factors, should be able to survive translation into a language appropriate for lower-class Americans when its initial development had been in a language appropriate for middle-class Americans.

Relevance of a 16 PF to Vocational Considerations

It seems likely that a brief discussion of the factors measured is in order at this point, so that the reader may judge for himself the probable utility of an instrument able to measure these factors. It is emphasized that such judgments of probable validity are no substitute for demonstrated results, but reasonableness is certainly a favorable consideration in evaluating the probability that a measuring instrument will yield useful information. (See Table 1.)

Table 1 is designed particularly for the reader who is not familiar with the 16 PF Tests. Nineteen factors are listed, sixteen of them being those that are represented in Forms A, B, and C of the 16 PF. Two factors (D and J) have never been adequately represented in the questionnaire medium among populations of normal adults, but they have appeared in ratings of adults and children, and it has been possible to construct questionnaire items to measure these two factors in children and adolescents. It was felt on theoretical grounds that either or both of these factors might be present in the population of culturally deprived or mentally retarded adults, and for this reason items designed to measure these factors were included. As will be seen below, the effort to measure these factors was successful.

The reader will notice that Factor K is not represented in any questionnaire test even though it has appeared very clearly in rating studies. No attempt was made to represent this factor in these studies since many previous attempts, both with children and with adults, have failed. It is entirely possible and meaningful, from a theoretical standpoint, for a factor to appear in the medium of ratings but not in the medium of questionnaires; this simply means that raters can see in the persons they rate a characteristic which the persons cannot describe in themselves.

The obverse of this situation is represented by the factors labeled Q_1, Q_2, Q_3, and Q_4. These are factors that have never been successfully found in ratings, but have appeared with strong predictive variance in questionnaires. The situation seems to be that these are things that people can describe about themselves but which a rater cannot see consistently enough or clearly enough to be able to rate precisely. This also presents no theoretical or practical difficulties.

The reader will note that Factor B, designated as intelligence or brightness, is not represented in the present study. The reason is that these items depend upon verbal understanding for their effectiveness, and a test designed to neutralize the effects of verbal understanding throughout the entire content would not normally be expected to reveal the influence of this factor. It was decided before the study was begun that no efforts to measure this particular factor would be made. For somewhat different reasons, it was decided that Factor Q_1 would be omitted from the present study. The items dealing with this generally

have been of a political or social opinion nature, and some attempts during the preliminary studies indicated that it would be difficult to represent these items with a culturally deprived or mentally retarded population without risking that the client would have difficulty understanding the item content. Therefore, Factor Q_1 was omitted from this attempt to build factored scales.

Finally, the reader's attention is called to each factor representation in various forms of the 16 PF and of related tests. It will be noted that some factors are missing in children but present in adolescents and adults, some are missing in children and adolescents but present in adults, and some are present in children and adolescents but missing in adults. This is not necessarily a lack of consistency between the various patterns. It is quite conceivable that something that is of very great importance in describing the behavior of children could disappear completely in terms of variation among adults. For example, it may be that children differ according to certain characteristics which are uniformly the same in adults, by virtue of the simple fact that training brings all adults to approximately the same level. It should be noted that this does not mean that the characteristic is missing, but only that no significant variance within that characteristic exists among adults.

Turning to the matter of the apparent meaningfulness of these patterns for vocational prediction, the reader is asked to make essentially his own judgments. It is not the purpose at this point to give empirical evidence, and the reader desiring significant empirical evidence should consult the Handbook or the various publications regarding the test. The

point is simply that the characteristics represented, assuming that they are measured with reasonable accuracy, should make possible a significant description of the characteristics of a person, and this description should be of significance not only for personal issues but for vocational ones.

Appropriateness of Research Setting

Granted that the development of a 16 PF for persons of limited literacy and verbal sophistication should be undertaken, there were substantial reasons why the setting in which the writer was engaged seemed an appropriate one for such development. The rehabilitation counselors in the area of Alabama served by the writer were, at that time and since then, quite active in service to mentally retarded persons, and the demographic characteristics of the population served in Alabama made available a large number of non-retarded but culturally limited clients. Thus, there would be no shortage of subjects for the experimental work.

A second consideration was the ready availability of "feed-back" data from rehabilitation centers which the writer served as consultant. In an evaluation center setting, where a client is evaluated eight hours a day, five days a week, by skilled persons, substantial information about such a client is developed which can serve as validating information for test data. Close relationships between the writer and the staffs of various rehabilitation centers, and substantial experience in joint analysis of failures of psychological prediction, assured that not only positive but also negative feed-back would be available.

Finally, the writer had developed a rather close relationship with

Raymond B. Cattell during the several months preceding initiation of this project, and the availability of Dr. Cattell as chief consultant meant that the work could benefit by consultation with the person who had done more to develop this theoretical approach than anyone else.

III. PROJECT METHOD

Factor Analysis

Multiple factor analysis is a rather technical and somewhat esoteric technique for precise statistical analysis according to a rather simple model. The assumptions made, the techniques used, and the criteria to be established are the subject of learned debate among experts in the field, but the basic idea is not so complex that it cannot be summarized.

The basic assumption is that if a large number of measurements are made upon a population of persons, there probably are in this total system fewer causes than measurements. Put another way, assuming for example that 100 different tests had been given to a large group of people, the chances are that the scores on these tests are the result of fewer than 100 different influences. It is likely that two tests, or perhaps a dozen tests in the system measure essentially the same thing. It is likely that a single cause contributes heavily to one test score, moderately to several others, mildly to several others still, and not at all to the rest. Based upon this assumption, the factor analyst attempts to find a reasonable number of influences that explain the responses of people to all the tests in the system, and then he attempts to identify these influences, which he calls "factors".

The 16 Personality Factor Test was built in exactly this manner. Taking a large number of questions about everyday life, with a substantial attempt to sample a wide variety of areas in which personality could be assumed to be involved, Cattell discovered that the responses of his subjects to these questions could be reasonably well accounted for by the assumption of 16 causes, hence 16 factors. There are substantial technical considerations of item selection, of scoring techniques, and of similar areas which need not concern us here. The point is that 16 influences upon personality are sufficient to account for a large proportion of the variation among people as they answer questions about themselves.

It should be emphasized that factor analytic method does not demand the types of random samples, representing a particular population, which characterize normative research. This is a consideration at a much later stage of development, when the factors have been identified, the test has been constructed, and the time for the development of precise norms has come. For the initial steps, the requirement is that the sample of persons tested be from widely differing backgrounds, of widely different characteristics, and generally representative of the full range of individual differences in various characteristics which are likely to exist in the population.

The Client Sample

Because of the emphasis upon various types of clients required by the factor analytic model, it was decided to obtain data from a variety of sources. Active vocational rehabilitation clients in the State of

Alabama, clients in vocational rehabilitation centers in Alabama and in other states, mentally retarded clients at the Sunland Training Center in Florida, mentally retarded clients at a rehabilitation center in Colorado, and finally mentally retarded clients in the institution for the retarded at Tuscaloosa, Alabama all contributed to the sample. The only requirement was that each client included in the sample had to be a person whose literacy level and verbal sophistication were such that existing tests at about the sixth-grade level would be beyond his understanding. Several samples were collected for various parts of the study.

The fact that clients referred to the writer by vocational rehabilitation counselors were included in the sample means that substantial variation in type of client resulted. Clients were referred directly by rehabilitation counselors, and were obtained also from rehabilitation centers for the blind, for the deaf, for severely disabled persons; included in the total sample, therefore, were clients of all types, those who were mentally retarded and those who were culturally deprived, those with other physical disabilities and those without such disabilities. The goal of having in the factor analysis a widely divergent sample of persons seems clearly to have been achieved.

The Experimental Items

As stated above, an analysis of this type begins with a large pool of tests, or items in this case, which are administered to a varied sample of persons. Two sets of items were designated in the initial work, items obtained by two different methods.

An experimental 16 PF Test, Form X, was obtained by liberal borrowing

of items from the Child Personality Questionnaire. The reason for this was simply that here were a group of items of known utility in measuring the factors, yet at the third-grade reading level. These items were not quite sufficient to constitute Form X, since some factors thought to be of relevance in adults do not appear significantly in children, and therefore items to measure these factors were not available. The writer, in collaboration with Cattell, developed items for these "missing" factors, so that finally a 160-item test, 10 items for each of the 16 proposed factors, was developed. This experimental test was designated the 16 PF, Form X.

The senior consultant, Raymond B. Cattell, on the basis of long experience in the development of tests of this type, proposed that it was quite likely that some items would be lost in the analysis, that is, some items would fail to preserve their characteristics when translated from a situation of measurement of children on the one hand to the measurement of intellectually and culturally limited adults on the other. It was decided, therefore, to construct a second set of 160 items, again 10 items for each of the 16 factors, and this set was designated the 16 PF, Form Y. Items were contributed by Cattell, by the writer, by William S. Wells, Ph.D., and by other psychologists and psychometrists. Discussion with rehabilitation specialists, particularly with evaluators in rehabilitation centers, yielded a substantial number of suggestions for items, some of which turned out very well. Forms A and B of the 16 PF, as well as the High School Personality Questionnaire, were examined for items that might permit translation to the simple language level of the

present requirement. The result was the 160-item set designated the 16 PF, Form Y.

Since Form X was a simple translation of an existing scale, and preliminary work with it indicated that, despite unknown validity, it could provide useful leads for rehabilitation, its routine use with persons who could not be tested with any other existing instrument was instituted in the writer's office and in the rehabilitation centers which the writer serves as consultant. Thus a large data pool regarding Form X built up quite quickly. Form Y data were somewhat more difficult to obtain, particularly since they would reach their maximum usefulness only in that case where both Form X and Form Y were available upon the same person at approximately the same time. Because of these considerations, several partial analyses of the data were undertaken, culminating finally in a complete analysis of a sample of 265 persons to whom both Form X and Form Y had been administered at approximately the same time.

Data Collection

Several methods of obtaining data were used in various samples participating in this project, and this is considered desirable because these various methods duplicate the techniques that are likely to obtain in actual daily use of any test. Clients were tested in small groups, particularly in those centers for retarded persons that participated. Clients were also tested individually, particularly in the case of those clients who were referred to the writer's office or who were at rehabilitation centers where the writer serves as consultant.

Some clients had the experimental test administered to them by being

asked to read the items and to respond by marking on an answer sheet. All such clients were those who could read at third-grade level or better (within the limitation that no clients who could read at sixth-grade level or better were included in the sample). Clients unable to read at third-grade level or better were tested either with oral administration, in which case the experimenter read the items and marked the client's answer sheet, or with tape recorded administration, in which case the items were presented by standard tape recording and the client marked the answer sheet himself.

Precise analysis of the differential effects of various types of administration is not properly conducted at this stage of research. Rather, such analysis belongs to a later stage when a final test has been developed, and precise norms have begun to be gathered. Approximate checks for factor patterns representing so-called "instrumental effects" were conducted with these samples, and no evidence of such problems occurred.

A final word about tape recorded methods of presenting the items to the clients concerns the obvious fact that the tape recorded presentation can be used only with those clients who are able to read sequential numbers on an answer sheet; virtually all clients with whom this method was tried, even severely mentally retarded clients, were able to do this. In those few cases where the client was unable to follow the answer sheet numbers, the test was presented orally with the examiner marking the answers.

A word is in order also about those very few clients who were

unable to participate meaningfully in the testing even under the conditions of the simplest presentation. It is not possible to be completely exact about the number of such persons, because the records are available only upon that group of clients who were referred to the writer's office or were seen in rehabilitation centers which the writer serves as consultant. Among these clients, which might be designated the writer's sample, less than 1% of those with whom tape recorded administration was attempted were unable to complete it. It is certainly true that this "writer's sample" represents a smaller proportion of institutionalized persons than other parts of the total sample, but it seems clearly indicated that the limitation placed upon the results and upon the use of the techniques that may be developed by inability of the client to handle tape recorded administration are very slight. Certainly the efficiency of the techniques with the vast majority of clients should make possible oral administration by a skilled examiner in those few cases where this is necessary.

IV. DATA ANALYSIS

The Oblique Simple Structure Criterion

Factor analysts generally recognize two methods of treating data, and these two methods are not in opposition, but rather are two different ways of dealing with the same types of data. The choice between which method shall be used is best made upon empirical and logical grounds. The criticism that a certain study should have used orthogonal rotation rather than oblique rotation, or vice versa, is not a valid criticism. Rather, it behooves the critic to re-analyze the data from the alternate

viewpoint, and to demonstrate that this analysis is as meaningful or more meaningful than the one used in a particular study. This point is made because any factor analytic set of data can be interpreted according to at least two major methods of analysis; the method chosen here is that of rotation to oblique simple structure. The reason for this choice is simply that the entire theoretical system upon which the 16 PF Tests are based uses the criterion of oblique simple structure, and it was definitely felt that the materials developed should be compatible with previous work in this area.

Leaving aside the mathematical technicalities, the choice becomes one between two approaches, which may be summarized as follows:

(1) The data can be analyzed in terms of the fewest possible number of factors, with the factors arranged in order of decreasing contribution to the variation in responses. This approach denies that there are underlying influences that may be represented mathematically, and states in effect that any mathematical model is as good as any other, provided that the predictions are the same. This so-called orthogonal model is the simplest one.

(2) Data may be analyzed in terms of the belief that there are underlying influences or causes which can be represented by the factors, and these causes may be related to each other. In this oblique model, a minimum number of factors is not the major criterion, but rather the overall simplicity and compellingness of the structure is taken as the criterion.

In the personality area, and particularly in the types of data

analyzed here, the first approach is likely to lead to the postulation of very few broad factors, each relatively difficult to interpret psychologically, but each designed so that its inclusion as a predictor will add the maximum amount of new information over that efficiency of prediction which was possible without it. The second, oblique model, on the other hand, leads to factors whose psychological nature is relatively easy to infer, and makes possible the use of substantial psychological insight both in the interpretation of the factors and in the application of the results to new situations. The second model is the one that was chosen for this data analysis.

Analysis by Item Packages

A factor analysis of this type can proceed with each item being considered a separate variable, or it can group the items in packages, so that each variable is a score composed of the sum of several item responses. Two advantages come to the investigator who uses the "package" approach. One of these is that fewer variables are to be analyzed, since, for example, 160 items can be represented as 32 packages of five items each, and so the matrix to be analyzed becomes one of 32 variables. Doing the same job by items requires the analysis of a matrix of 160 variables, and at the time that this study began, no techniques were available for analysis of a matrix larger than 120 variables, even by the use of the largest high-speed computers.

The second advantage is that item package scores are inherently more stable than single item scores. Small chance errors cancel themselves out when one deals with a score which is the sum of several single items.

While these two advantages seem quite substantial, there are disadvantages also. In the last analysis, one is not really concerned about item packages but rather about the factor content of single items, and the packages must be "unwrapped" at a late stage in the analysis in order to make this development possible. The main reason for working with item packages has been the limitation upon computing facilities, and during the course of this research, the writer developed techniques making possible the direct analysis of single items in extremely large groups. Thus, while part of the data were analyzed by the older, simpler methods, the final analysis was conducted with the newer and more satisfactory methods. In fact, the final complete analysis of each of the 320 experimental test items was handled in a single giant matrix quite late in the course of the project.

A word may be in order about the substantial difficulties that arose during the early analyses, since these difficulties led to the felt need for techniques capable of handling the large matrices of single items. The problem is that one can easily analyze the packages, but then some precision is lost in the "unwrapping" of the packages, and in fact some variance is lost throughout the technique. A second stage of the project used the approach of handling single items, which by that time seemed to be more satisfactory, but then computing limitations required the handling of matrices no larger than 80 variables. This in turn means that so-called "marker variables" have to be carried from one analysis to the next, a technique whereby variables known to define certain factors are brought into each subsequent analysis along with some new ones, so that

we may be certain that the new factors are the same as the old ones. Each time one does this, one loses a certain amount of the precision that is desirable. While this technique was used with the original 16 PF development (because no other choice existed at that time), it seemed very clear as the analysis developed that ideal handling of the data would be best accomplished by computing techniques that enabled the writer to look at the entire picture in one matrix.

The writer therefore set about the business of developing computing programs that would transcend the limitations of computer storage space. There is no inherent reason why programs handling matrices as large as one desires cannot be built, but it had not been done before simply because the pressing need to do this had not occurred to anyone who had the appropriate training to develop the programs. Unfortunately, involvement in psychometric research and computer programming skills do not coexist in large numbers of psychologists. Once the writer decided that programs of unlimited size were needed, it was a relatively straightforward matter to write these programs, and they are now available. In fact, it may be said that the making available of large-scale programs of this type is a significant by-product of this project.

V. RESULTS

Item Packages from Form X

Responses to Form X given by 325 persons were punched into IBM cards by the Data Center of the State Department of Education, Tallahassee, Florida, under the supervision of Dr. Bob Gates. Complete programs for handling the data in the desired manner were not available in that

department, or in the IBM 709 center at Florida State University where additional computing was to be done. However, after substantial reprogramming was conducted by collaborative effort of programmers at Florida State University, Cattell, and the writer, the development of the correlation matrix and the factor extraction were possible. Thirty-two packages were involved, each composed of five items from Form X. The packages were made up by taking five items from Form X that were known to involve a certain factor, and since 10 items for each factor were included in Form X, this made two packages per factor or a total of 32 packages.

Substantial negotiations were conducted to make appropriate programs available at Florida State University, but the writer discovered at about this point that it is one thing to have a program working in a certain laboratory and quite something else to transfer it to a different laboratory. Thus, while we discovered that the programs we needed were supposedly available at certain laboratories in the United States, we also discovered that it was extremely difficult to obtain these programs for use at Florida State University. This was not a matter of lack of cooperation or lack of communication, but simply a reflection of the state of the computer art and technology at that time. If the writer may be permitted a side remark, things have changed considerably during the three years of this project.

Finally, when it became evident that the computing could not be done at Florida State University because programs simply were not available, arrangements were made to shift to the University of Illinois facility

where Cattell's programs were installed and running. The analysis was done under Cattell's supervision, and the results were quite encouraging.

In the first place, 16 clear factors appeared, as had been hypothesized. It was not possible to say that a larger number of factors might be present, but certainly 16 clear factors, as hypothesized, emerged from the analysis. Rotation by the rotoplot technique (Cattell & Foster, 1963), a combination of computer oriented and visually directed methods, produced a quite clean simple structure after seven rotation cycles, with an overall hyperplane count of 68% within a ± .10 hyperplane range.* In this structure, 23 of the 32 packages appeared in the predicted position, four packages appeared in logical positions other than the predicted ones by virtue of the confusion of two factors with each other, and only five packages failed to find their hypothesized place. A review of the data at this point in the analysis revealed also that some errors had been made in communicating the requirements for keypunching, so that the data were

---

*A note of explanation for the non-technical reader is in order. The concept of simple structure is that, in any system of causes and effects, any one effect is probably the result of fewer than the total number of causes in the system. "Percent hyperplane count" shows the degree of simplicity of the structure. In this case, with 16 factors (causes), a 68% hyperplane count means that the average item is influenced by only 5 of the 16 causes, and unaffected by the remaining 11 (68%). The "± .10" range is the range of values considered as essentially zero. In the type of problem discussed here, hyperplane counts above 55% are considered highly definitive of simple structure.

somewhat contaminated and somewhat lacking in proper reliability. The errors were not sufficiently large to invalidate the analysis, but they did suggest that further attempts to purify this particular analysis should be abandoned.

Thus, at the point described, working with somewhat imperfect data, reasonably good general confirmation of the factor structure had been obtained, although the confirmation was by no means sufficient to warrant final confidence.

Analysis of Single Items from Form X

Because the difficulties inherent in the analysis of packages had already begun to appear at this stage, it was decided that the analysis would now shift to the handling of single items. The writer had by this time developed the appropriate programs for conducting the analysis on an IBM 1620 computer which was readily available to him (between the hours of midnight and 4 a.m.). Consequently, an intercorrelation matrix representing the first 80 items of Form X, upon the same basic population of 325 clients, was constructed. This matrix was factor analyzed by the centroid method (Thurstone, 1947) and 16 factors with roots greater than unity emerged. These factors were rotated by the rotoplot method, (Cattell & Foster, 1963) and a compelling simple structure emerged after 29 rotations (57% count within a ± .10 hyperplane).

This was clearly a significant result, and it should be noted that most of the items behaved as predicted, showing variance in the predicted factors. Perfect correspondence between previous knowledge regarding these items and the results of this analysis was not expected, since the

population was quite different, and some items can be expected to change their factor composition slightly when moving to a radically different population. The point to be made is that the 16 factors were clearly defined, that the simple structure was of the order of simplicity that was anticipated, and that the item content for each factor made excellent sense. Thus, the analysis clearly confirmed the existence of the personality factors when these items were applied to the relevant population, and there were no errors in these data. It was felt that the basic question of whether this type of analysis could yield meaningful results had been answered affirmatively.

The procedure of analyzing single items was then continued with a new sample of questions, but with the same sample of clients. Thirty-two marker variables, two for each of the 16 factors, were selected out of the first 80 items which had just been analyzed, and these 32 marker items were combined with 48 additional items from Form X, the last 48 not previously analyzed. The resulting 80 items were intercorrelated, the correlation matrix was factored, and rotation was begun.

Marker variables make rotation somewhat easier. After 11 rotation cycles, a clear simple structure emerged with these new data (60% count within a $\pm$ .10 hyperplane), and the markers had generally retained their position so as to make the identification of the factors quite clear. Again, with this quite compelling simple structure, the item content was sensible and consistent with the factor meaning, and thus further confirmation of the existence of the factors in these items and of the basic structure of the data appeared.

The process of choosing markers and combining these with new items, refactoring, and then rotating to simple structure was continued with yet a third set. Again, this time after 15 rotation cycles, a compelling simple structure emerged (58% count within a ± .10 hyperplane). Again the item content was stable and sensible, and again the factor definitions appeared quite clear.

Thus, at this stage of the analysis, it had been demonstrated that the factors appeared clearly among the Form X items, that rotation of factors derived from single items was possible, that normal standards of simple structure definition could be applied successfully, and that markers carried over from one set of items to the next maintained their factor position as anticipated.

Analysis by New Form X Packages

At this point, it was decided to attempt to develop more stable markers for subsequent analyses than the markers furnished by single items. The technique used was to combine Form X items which had previously shown clear factor patterns into appropriate factor packages. Thirty-two packages were created, 16 factors each being represented by two packages. These 32 packages were intercorrelated, over the same population, centroid factors were extracted, and rotation was begun.

After eight rotations, a compelling simple structure emerged (69% count within a ± .10 hyperplane). Twenty-nine of the 32 factor packages appeared in the structure exactly at the predicted place and with the predicted values; three packages failed to show the proper loadings, but each of these three was only one of a pair defining that particular

factor, so that each factor was represented by at least one appropriate package in the final structure and 13 of the 16 factors were represented by both of the appropriate packages. It was concluded at this point that the structure was essentially what was expected and desired, and that the analysis was yielding the appropriate results.

Analysis of Form X Packages Combined with Form Y Items

A matrix was now constructed upon a new population, this time a sample of 216 clients who had taken both Form X and Form Y at approximately the same time. In no case had more than one week elapsed between the time the client took one form and the time he took the other. The Form X package scores were defined, and these 32 scores were combined with 48 items from Form Y, producing again an 80-item matrix. Intercorrelations were computed and 16 centroid factors were extracted. Rotation was conducted and a compelling simple structure emerged (60% count within a ± .10 hyperplane). Again, the item content and the package content made sense in terms of each of the factors, and it was felt that the analysis was continuing appropriately.

The above process was repeated with the marker items from Form X and a second set of 48 items from Form Y. Again a reasonably compelling simple structure was obtained, although not quite so precise a structure as had previously been the case. Again, item content and package content made sense when analyzed in terms of factor meanings.

Difficulties Inherent in the Package Marker Approach

A sufficient number of analyses had been conducted by this time to permit a judgment as to the eventual efficacy of this technique. The results

caused some concern. It was found that the adequacy of definition of the structures had been decreasing slightly but significantly with each time that marker variables from the immediate previous study were carried into a new one. Marker variables that had been used in all the analyses lost some of their efficiency as subsequent steps were taken. This is by no means surprising, since it is in fact this characteristic that would be expected from the relative instability of single-item scores or even scores composed of the sums of responses to three- or four-item packages. The point is that this method of tackling the problem a little bit at a time, and extending into each new subset of variables a reasonably good definition of the factors, inevitably results in gradually less adequate definition of the factors, until the investigator has difficulty being certain that significant changes in factor definition have not crept in. This type of result is completely characteristic of this method, and in the past it was accepted simply as an inevitable thing, since no computing methods were available to tackle the problem in any other manner. However, the writer in this case was confronted in about the summer of 1964 with a choice that had not previously been available.

The writer discovered, by personal communication from a friend and colleague, that arrangements might be made to permit the analysis of the data at an Army Computation Center charging an extremely favorable rate for an IBM 7094 computer. The problem would be to develop the programs appropriate for the use of this computer, but the large size and extremely high speed of this machine made possible the consideration of programs that had not previously been used anywhere. The writer thus faced the

choice between proceeding as he had been, with all the imperfections admittedly inherent in this "step-by-step" approach, or of taking time to develop the appropriate programs and tackling the problem at one single step with the large machine. There was no precedent for the attempt to handle this as one large problem, but there appeared to be no theoretical difficulties, and the writer became convinced, after consultation with Cattell and with others, that the problems inherent in large-scale analysis could be solved. Additional factors favoring the decision to develop large programs included the fact that a new project was beginning in the writer's setting, a project which would benefit tremendously from the availability of the new techniques.

Attempts were therefore made to develop appropriate programs capable of handling the entire matrix of 320 items (160 from Form X and 160 from Form Y) at a single step, and these efforts were successful by late fall of 1964.

The Large Matrix

Intercorrelations of each of the 320 items with each of the others (a matrix of more than 50,000 entries) were calculated and 16 factors were extracted by the centroid method. Rotation was conducted by the newly developed Maxplane method (Eber, 1965). The final results confirmed the correctness of the decision to attempt development of the new, large-scale techniques. The rotation yielded a compelling simple structure with 62.3% count within a $\pm$ .10 hyperplane. The factors were clearly interpretable, all item content being sensible, and this appeared finally to be the desired result.

Appendix A contains a listing of the factors included in the study, a description of each in relatively non-technical terms but in somewhat more detail than in Table 1, and a listing of two sample items selected from those items found to define the factor. Complete data of the large analysis may be obtained from the writer by seriously interested researchers; the volume of the data makes presentation in any other form impractical.

VI. DISCUSSION

The above results imply firm, positive answers to the four questions defined as areas of concern in Section II. It seems rather clear that items yielding meaningful self-descriptions can be constructed within the limitations of simplicity and low literacy, that these items yield personality description patterns that match known ones, that this "match" is sufficiently good to permit the transfer of previous knowledge to the use of these new patterns, and finally that these patterns are sufficiently stable to form the basis of further research. Some discussion of each of these points seems in order.

Eight factor analyses of several different types, of several sizes, and using in some cases item packages and in other cases single items all yielded clearly defined simple structure patterns of the type anticipated. In each case, the structure was not only statistically compelling but psychologically meaningful. The anticipated factors were clearly found, defined in terms of item content that had been anticipated. This does not mean that every single item behaved in the total analysis exactly as predicted, since such a thing is never to be expected in any large-scale

analysis of this type. However, when eight overlapping factor analyses each yield clearly defined the 16 factors that had been anticipated, it seems clear that the simplicity of the items and the low literacy level required to answer them meaningfully do not constitute barriers to meaningful self-description.

Additional confirmation of the validity of the results comes from the fact that the factors match known ones. This is after all what is to be expected, since the vast body of theoretical research in human personality using this particular approach has repeatedly found these same factors represented in various types of tests. Still, expected or not, the fact that no difficulty whatever was experienced in matching the new-found patterns with previously defined ones gives added confidence in the validity of these results.

The literature contains a vast number of studies relating the personality factors to various criteria. It is not properly cautious to use these data immediately with the factors as measured by the techniques developed here, but certainly the process of validation can begin without having to start from scratch. Our knowledge that, for example, certain occupations are best performed by persons who are quite "extraverted" does not depend upon the particular test which was used to define extraversion. Rather, to the degree that this characteristic can be adequately defined, the previous knowledge regarding it can be applied to the particular case in point. Since extraversion is a composite of several 16 PF factors, and since these same factors appear in the simple items used in this study, it becomes immediately possible to utilize our knowledge

regarding this characteristic in interpreting results obtained with any client using the new techniques. The same thing is true regarding other second-order factors besides extraversion, such as anxiety, independence, and alert-poise; similarly, the same is true for other composite scores defining such areas as various vocational potentials, academic potential, neurosis, etc. It is to be emphasized that uncritical acceptance of previous findings and the transfer of this knowledge into work with the simpler instruments that will arise from this study is not recommended; however, the previous knowledge presents a vast body of hypotheses to be tested by further research, and the investigation of the predictive validities of test instruments based upon the present work will not need to begin in isolation.

The advantage that comes from the fact that this study fits into a large body of personality theory cannot be overemphasized. Since the relationships between the personality factors and various criterion situations have been the subject of extensive previous research, it now becomes a matter of constructing instruments to measure the factors as precisely as possible within the limitations of item simplicity and low reading level, and then the predictive validity of the new test materials can be directly calculated upon the basis of the degree to which they measure the relevant factors.

Finally, the present results, and the personality test materials that can and will be derived from them, make possible the application of personality factor theory to a large number of human situations where previous application of these principles was not possible. For the first

time since the approach represented by the 16 PF Tests was developed, the personality factors can be measured upon a total population, one unrestricted as to reading competence, intellectual and educational competence, and socio-economic background. A large number of applied situations focusing upon the needs of mentally retarded or culturally deprived adults demand the attention of the psychologist; the present study helps to remove barriers standing in the way of the application of psychological techniques to the service of the less advantaged persons in the population.

Table 1

Factor	Brief Description	Represented Previously in:*							Included in Project
A	Outgoing	A	B	C	HS	Ch	ES	R	yes
B	Bright, alert	A	B	C	HS	Ch	ES	R	no
C	Emotionally stable	A	B	C	HS	Ch	ES	R	yes
D	Impulsive				HS	Ch	ES	R	yes
E	Assertive	A	B	C	HS	Ch	ES	R	yes
F	Enthusiastic	A	B	C	HS	Ch	ES	R	yes
G	Conscientious	A	B	C	HS	Ch	ES	R	yes
H	Venturesome	A	B	C	HS	Ch	ES	R	yes
I	Tender-minded	A	B	C	HS	Ch	ES	R	yes
J	Aloof, arrogant				HS	Ch	ES	R	yes
K	Conforming to culture							R	no
L	Suspicious	A	B	C				R	yes
M	Imaginative	A	B	C				R	yes
N	Shrewd	A	B	C		Ch	ES	R	yes
O	Apprehensive	A	B	C	HS	Ch	ES	R	yes
Q_1	Experimenting	A	B	C					no
Q_2	Self-sufficient	A	B	C	HS				yes
Q_3	Self-disciplined	A	B	C	HS	Ch			yes
Q_4	Tense	A	B	C	HS	Ch	ES		yes

*A, B, C = 16 PF, Forms A, B, C respectively
HS = High School Personality Questionnaire
Ch = Child Personality Questionnaire
ES = Early School Personality Questionnaire
R = Rating Data

References

Cattell, R. B. Personality and motivation structure and measurement. Yonkers, N. Y.: World Book Co., 1957.

Cattell, R. B., & Eber, H. W. Handbook for the 16 personality factor test. Champaign: IPAT, 1964.

Cattell, R. B., & Foster, M. J. The rotoplot program for multiple single-plane, visually-guided rotation. Behav. Sci., 1963, 8, 156-165.

Cattell, R. B., & Scheier, I. H. The meaning and measurement of neuroticism and anxiety. New York: Ronald Press, 1961.

Eber, H. W. Toward oblique simple structure: maxplane. Multivar. Behav. Res., in press.

Research Division, MacDonald Training Center. Predicting vocational capacity of retarded young adults. Tampa, Fla.: MacDonald Training Center Foundation, 1963.

Thurstone, L. L. Multiple factor analysis. Chicago: Univer. Chicago Press, 1947.

Final Report
VRA Grant No. RD-1059p-62

Appendix A

The following is a list of the factors found to be represented in the low-literate items, as well as a listing of some sample items for each factor.

Factor A

At the positive end, this factor represents a person who is outgoing in interpersonal relationships; at the negative end, there is social reserve. Here, as is the case with most of the factors, value judgments regarding which type of score is "better" are not justified; a more meaningful question would be "better for what purpose?" Sample items and loadings:

*Would you like to start a new club	or	don't you like clubs	.446
Do you like to be active in social things	or	would you rather be alone	.578

Factor C

The high score on this factor describes an emotionally stable person; the low score describes an overemotional person, likely to "come to pieces" under emotional stress and lacking in self-confidence. Sample items and loadings:

Can you do most things well	or	can others do things better	.369
In your group is someone else the leader	or	are you the leader	.380

Factor D

A high score on this factor describes an impulsive person, poorly disciplined; low scores describe a more mature person able to keep impulses in check. Sample items and loadings:

When a small thing upsets you, do you get so mad you want to throw things	or	can you keep calm	.474
Do you talk back to your parents	or	are you afraid to	.333

---

*The underlined response is factor positive.

Factor E

A high score on this factor describes an assertive, dominant person, while a low score describes a passive, submissive one. Sample items and loadings:

Would you rather make sure people like you	or	make sure they do what you want	.351
Would you rather listen to a teacher	or	talk yourself	.353

Factor F

A high score on this factor describes an enthusiastic, surgent person; a low score describes a sober, more quiet person. Sample items and loadings:

In a game, would you rather just play around	or	work hard to win	.439
Should everyone have an airplane	or	are cars enough	.308

Factor G

A high score describes a conscientious moralistic person; a low score describes an expedient person. Sample items and loadings:

When people ask for help on a test do you let them do their own work	or	help them if nobody's watching	.351
When you write something and it looks messy do you copy it over	or	leave it as it is	.436

Factor H

A high score describes a venturesome, stress resistant person; a low score describes a person who can function best in stable and predictable situations. Sample items and loadings:

When a visitor comes to your house do you talk to him first	or	do you feel too shy	.483
Would you rather go to a school	or	go on a long car trip	.416

Factor I

The high scoring person is described as tender-minded, sensitive, and probably overprotected; the low scoring person is tough-minded and insensitive. Sample items and loadings:

In a play would you rather be a jet pilot	or	a famous writer	.440
Would you rather be a good musician	or	a good soldier	.335

Factor J

The high scoring person is described as aloof, somewhat arrogant, and oriented to precise work; the low scoring person puts peer group relationships before precision of workmanship. Sample items and loadings:

Do you feel sorry for criminals	or	do they deserve to be punished	.476
Which is more important, who you know	or	what you know	.344

Factor L

The high scoring person is described as suspicious; the low scoring one is a trusting person. Sample items and loadings:

Were your parents always fair to you	or	were they often unfair	.411
Do most people really like to help a person	or	do they just pretend to	.362

Factor M

The high scoring person is described as imaginative and phantasy-oriented; the low scoring person is precise and environment-oriented. Sample items and loadings:

Do you think the world needs more people who do things just right	or	more dreamers	.317
When someone talks too much does your mind wander	or	do you still listen to everything he says	.341

Factor N

The high scoring person is described as shrewd, cynical, sophisticated; the low scoring person is naive and forthright. Sample items and loadings:

Would you rather read funny books	or	do arithmetic	.400
Will anyone tell a lie to keep out of trouble	or	will most people tell the truth anyway	.422

Factor O

The high scoring person is described as apprehensive and ready to feel guilty; the low scoring person is comfortable and secure. Sample items and loadings:

When your mother is angry with you do you feel all right anyway	or	do you feel like crying	.361
If you do something wrong do you worry about it a lot	or	soon forget it	.342

Factor Q_2

A high score on this factor describes a self-sufficient person, while a low score describes one whose orientation to the group makes him somewhat dependent upon group opinion. Sample items and loadings:

Do you follow your own idea of what is best	or	do the same as other people	.300
When you need to find your own way in a strange place is it better to use a map	or	to ask somebody	.310

Factor Q_3

A high score describes a self-disciplined person, while a low score describes a lax, rather unreliable person. Sample items and loadings:

Is it hard to change your ways	or	easy	.373
Can you usually find your way in a strange place	or	do you get lost	.327

Factor Q_4

A high score describes a tense person, while a low score describes a comfortable, relaxed person. Sample items and loadings:

When a test is over do you just forget about it	or	keep worrying about it	.399
Are more wrecks caused by people driving too fast	or	by blocking the road driving too slow	.323

It should be emphasized that the above items are just samples, prototypes. It should not be assumed that these items are pure in the particular factor, that is, that they do not have loadings in another factor as well. Such "factor pure" items are quite rare. Nevertheless, it is felt that the above described items give something of the sense of each factor beyond the simple description that is listed. For further detailed descriptions of the factors, attention is called to the work of Cattell (1957).